

Liver cirrhosis prediction for patients with Wilson disease based on machine learning: a case–control study from southwest China

Ke Chen^{a,b,*}, Yang Wan^{a,b,*}, Ju Mao^a, Yuqing Lai^a, Gesang Zhuo-ma^a and Peiwei Hong^{a,b}

Objectives Wilson disease (WD) is a rare autosomal recessive disease caused by an *ATP7B* gene mutation. Liver cirrhosis is an important issue that affects the clinical management and prognosis of WD patients. Blood routine examination is a potential biomarker for predicting the occurrence of liver cirrhosis in WD. We aim to construct a predictive model for the occurrence of liver cirrhosis using general clinical information, blood routine examination, urine copper, and serum ceruloplasmin through a machine learning approach.

Methods Case–control study of WD patients admitted to West China Fourth Hospital between 2005 and 2020. Patients with a score of at least four in scoring system of WD were enrolled. A machine learning model was constructed by EmpowerStats software according to the general clinical data, blood routine examination, 24 h urinary copper, and serum ceruloplasmin.

Results This study analyzed 346 WD patients, of which 246 were without liver cirrhosis. And we found platelet large cell count (P-LCC), red cell distribution width CV (RDW-CV), serum ceruloplasmin, age at diagnosis, and mean corpuscular volume (MCV) were the top five important predictors. Moreover, the model was of high accuracy, with an area under the receiver operating characteristic curve of 0.9998 in the training set and 0.7873 in the testing set.

Conclusions In conclusion, the predictive model for predicting liver cirrhosis in WD, constructed by machine learning, had a higher accuracy. And the most important indices in the predictive model were P-LCC, RDW-CV, serum ceruloplasmin, age at diagnosis, and MCV. *Eur J Gastroenterol Hepatol* 34: 1067–1073

Copyright © 2022 The Author(s). Published by Wolters Kluwer Health, Inc.

Introduction

Wilson disease (WD), often referred to as hepatolenticular degeneration, is a rare autosomal recessive disease caused by an *ATP7B* gene mutation [1,2]. Then, *ATP7B* gene defects can lead to dysfunction of copper-transporting ATPase 2, which results in copper deposition in the liver, brain, cornea, and other organs [3]. Its prevalence rate ranges from 2.7 to 22 per million in different regions, which might be underestimated because it was calculated based on symptomatic patients [1]. But the prevalence rate confirmed by genetic analysis ranges from 30 to 885 per million in different populations [1].

Overall, 40–50% of WD will manifest as symptomatic liver disease [4]. But it may manifest in a broad spectrum

of hepatic diseases, from mild hepatic dysfunction to liver failure [4]. Advanced fibrosis is an important determinant that has an influence on the prognosis of individuals with chronic active hepatitis. Patients who develop cirrhosis and portal hypertension as a result of fibrosis advancement are caused by persistent active inflammation [4]. Liver cirrhosis is an important issue that affects the clinical management and prognosis of WD patients [4]. There is a model of end-stage liver disease (MELD) score, which is calculated by utilizing total bilirubin, creatinine, and international normalized ratio, that can evaluate the degree of illness and mortality of WD patients [4,5]. The Nazer score is used to predict the mortality of WD patients without liver transplantation [5]. However, there is no model that can be used to predict the development of liver cirrhosis in patients with WD.

Blood-based erythrocyte, leukocyte, and platelet parameters are an inexpensive method and easy accessibility of measurement, which is a potential biomarker for some medical conditions [6]. Machine learning is an algorithm of techniques for learning and generalizing from data. It has the potential to improve human behavior and event predictions [7,8]. To the best of our knowledge, machine learning has not been applied to predict the occurrence of liver cirrhosis in patients with WD. Here, the study aims to find the most important indicators and construct a predictive model for the occurrence of liver cirrhosis in patients with WD using general clinical information, blood routine examination, urine copper, and serum ceruloplasmin through a machine learning approach, aiming to provide clinicians with a tool for clinical decision-making.

European Journal of Gastroenterology & Hepatology 2022, 34:1067–1073

Keywords: blood routine examination, hepatolenticular degeneration, liver cirrhosis, machine learning, Wilson disease

^aDepartment of Geriatric Medicine and Neurology, West China School of Public Health and West China Fourth Hospital and ^bWest China-PUMC C.C. Chen Institute of Health, Sichuan University, Chengdu, People's Republic of China

Correspondence to Peiwei Hong, MD, Department of Geriatric Medicine and Neurology, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, 610041, People's Republic of China
E-mail: hongpeiwei1988@qq.com

*Ke Chen and Yang Wan contributed equally to the writing of this article.

Received 18 May 2022 **Accepted** 29 June 2022

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Methods

Patients

The term 'WD' was utilized to search the front pages of medical records of patients with WD, which each patient had an independent medical record number for each hospitalization in medical record system of West China Fourth Hospital, Sichuan University. The search period is limited to January 2005 to May 2020. The clinical data associated with the scoring system of WD were acquired [9]. Moreover, two experienced neurologists evaluated the total score of every patient, and patients with a score of at least four were enrolled, which defined in guideline that diagnosis with WD was established [9].

Patient and public involvement

This study was approved by West China Fourth Hospital Ethics Committee (Approval No.: HXSY-EC-2020088) and was performed in line with the Declaration of Helsinki. No patient was involved in the design, recruitment, conduct, or dissemination plans of this research.

Study design and data collection

The medical records of patients enrolled were reviewed retrospectively. The real-life information of patients from the first hospitalization in our hospital was collected from electronic medical records, such as epidemiological, demographic, medication history, clinical, laboratory, treatment, and discharge diagnosis data.

Outcome

The primary outcome of the study is the development of liver cirrhosis. Patients enrolled were divided into two groups based on their liver cirrhosis disease status: the without liver cirrhosis (WLC) group, and the liver cirrhosis group. Cirrhosis was diagnosed using ultrasonographic liver features according to established criteria [10].

Predictor variables

Gender, age, age at onset, age at diagnosis, delayed diagnosis time, course of the disease, usage of drug [D-penicillamine, zinc, and dimercaptosuccinic acid (DMSA) or dimercaptosulphonate sodium (DMPS)], blood routine examination, 24 h urinary copper, and serum ceruloplasmin were utilized as predictors. The missing value of continuous variables (including blood routine examination, 24 h urinary copper, and serum ceruloplasmin), were filled by mean.

Statistics

All statistical analyses were conducted using EmpowerStats software (X&Y Solutions, Inc., Boston, Massachusetts, USA). The continuous variable data were expressed as mean \pm SD, and the difference between groups was examined using *t*-test. Meanwhile, categorical variable data were expressed as numbers and percentages, of which the difference between the groups was analyzed using Chi-squared test or Fisher's exact test.

Predictive model development

The R packages with eXtreme Gradient Boosting (XGBoost), XGBoostExplainer, and machine learning in R were utilized for constructing the predictive models. Meanwhile, the R code is implemented in EmpowerStats. The eligible patients with WD were enrolled and randomly partitioned into training sets (75%) and testing sets (25%) for the algorithm [7,11,12]. Gini impurity was utilized to measure a variable's importance, which presented the contribution of the predictor variables in the predictive model [7,11]. And the relative importance score was utilized to present the importance of predictor variables. The most important predictor variable was set as 1, and the score of other variables is the ratio of the variable to the important variable [7,11]. A receiver operating characteristic curve (ROC) was utilized to assess the predictive model [7,11]. And the area under the ROC (AUC) was utilized to quantify the predictive model [7,11].

Results

Neurologists analyzed 1498 medical records of 380 patients. Following that, 346 patients had at least four according to the scoring system of WD. Of the 346 patients,

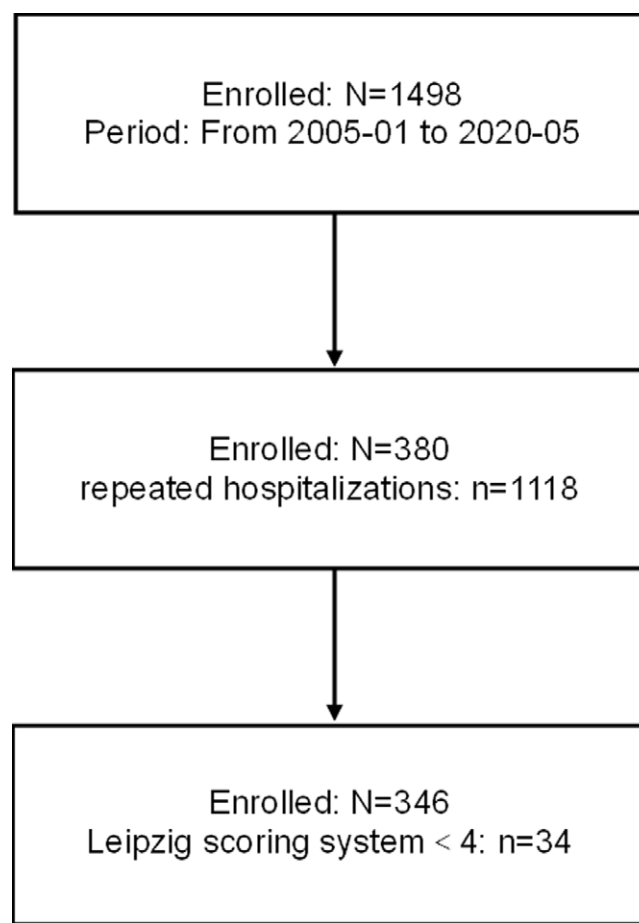


Fig. 1. Medical records screening procedure. One thousand four hundred ninety-eight medical records were acquired from department of medical record, West China Fourth Hospital, Sichuan University. The 1498 medical records included 380 independent Wilson disease patients. Among them, there were 346 patients who were of Leipzig scoring system no less than 4, and were enrolled for data analysis.

246 patients were WLC, and the remaining 100 patients had liver cirrhosis. The procedure of patient screening is summarized in Fig. 1.

There were no differences in gender, delayed diagnosis time, course of disease, usage of D-penicillamine, usage of zinc, and usage of DMSA/DMPS between the WLC group and the liver cirrhosis group. When compared to the liver cirrhosis group, WLC group had lower age (21.74 vs. 25.28, $P=0.021$), age at onset (17.44 vs. 21.77, $P<0.001$), and age at diagnosis (18.65 vs. 22.92, $P=0.002$). White blood cell count (4.94 vs. 4.56, $P<0.001$), and neutrophil count (2.84 vs. 2.68, $P=0.008$) were higher in the WLC group than in the liver cirrhosis group. The WLC group had a lower number in lymphocyte count (1.88 vs. 2.25, $P=0.004$) than the liver cirrhosis group. And there were no differences in monocyte count (0.44 vs. 0.40, $P=0.122$), eosinophil count (0.17 vs. 0.12, $P=0.06$), basophil count (0.02 vs. 0.02, $P=0.331$), neutrophil to lymphocyte ratio (NLR, 2.15 vs. 2.24, $P=0.447$), neutrophil percentage (0.58 vs. 0.58, $P=0.615$), lymphocyte percentage (0.37 vs. 0.53, $P=0.438$), monocyte percentage (0.09 vs. 0.09, $P=0.121$), eosinophil percentage (0.03 vs. 0.03, $P=0.654$), and basophil percentage (0.00 vs. 0.00, $P=0.338$) between these two groups. Meanwhile, the WLC group had a higher

value in red blood cell count (4.56 vs. 4.06, $P<0.001$), hemoglobin (129.34 vs. 116.12, $P<0.001$), and hematocrit (HCT, 38.86 vs. 35.13, $P<0.001$) than the liver cirrhosis group. The WLC group, on the other hand, had a lower value in mean corpuscular volume (MCV, 86.48 vs. 88.10, $P=0.026$), red cell distribution width CV (RDW-CV, 13.88 vs. 15.59, $P<0.001$), and red cell distribution width SD (RDW-SD, 46.28 vs. 49.86, $P<0.001$). And there were no differences in mean corpuscular hemoglobin (MCH, 28.68 vs. 28.96, $P=0.261$), and mean corpuscular hemoglobin concentration (MCHC, 331.46 vs. 329.28, $P=0.077$). In addition, the WLC group had a higher value in platelet count (139.34 vs. 87.04, $P<0.001$), mean platelet volume (MPV, 11.54 vs. 11.16, $P=0.009$), platelet distribution width (PDW, 17.05 vs. 15.52, $P=0.013$), plateletcrit (PCT, 0.17 vs. 0.12, $P<0.001$), and platelet large cell count (P-LCC, 49.88 vs. 30.83, $P<0.001$), which is defined as number of platelets with a volume of more than 12 fL. However, there was no difference in platelet large cell ratio (P-LCR, 37.78 vs. 36.47, $P=0.094$). Meanwhile, the WLC group had a lower value in urinary copper (3.72 vs. 4.55, $P=0.003$), and serum ceruloplasmin (52.48 vs. 56.94, $P=0.039$). The detailed information and units of all variables are shown in Table 1.

Table 1. Characteristic of Wilson disease patient enrolled

	WD without Liver Cirrhosis	WD with Liver Cirrhosis	<i>P</i>
<i>n</i>	246	100	
Gender			
Male	139	52	0.445
Female	107	48	
Age (year)	21.74±9.62	25.28±11.60	0.021
Age at onset (year)	17.44±8.87	21.77±10.72	<0.001
Age at diagnosis (year)	18.65±9.02	22.92±10.99	0.002
Delayed diagnosis time (month)	14.10±20.09	19.03±31.62	0.728
Course of disease (year)	4.33±6.03	3.56±5.84	0.068
Usage of drug			
D-penicillamine	136 (55.28%)	51 (51.00%)	0.468
Zinc	43 (17.48%)	15 (15.00%)	0.576
DMSA/DMPS	15 (6.10%)	2 (2.00%)	0.11
White blood cell count (10 ⁹ /L)	4.94±1.57	4.56±2.51	<0.001
Neutrophil count (10 ⁹ /L)	2.84±1.15	2.68±1.89	0.008
Lymphocyte count (10 ⁹ /L)	1.88±2.89	2.25±5.88	0.004
Monocyte count (10 ⁹ /L)	0.44±0.88	0.40±0.55	0.122
Eosinophil count (10 ⁹ /L)	0.17±0.31	0.12±0.11	0.06
Basophil count (10 ⁹ /L)	0.02±0.07	0.02±0.02	0.331
NLR	2.15±2.05	2.24±2.27	0.447
Neutrophil percentage	0.58±0.13	0.58±0.12	0.615
Lymphocyte percentage	0.37±0.42	0.53±1.46	0.438
Monocyte percentage	0.09±0.15	0.09±0.10	0.121
Eosinophil percentage	0.03±0.06	0.03±0.02	0.654
Basophil percentage	0.00±0.01	0.00±0.01	0.338
Red blood cell count (10 ¹² /L)	4.56±0.58	4.06±0.86	<0.001
Hemoglobin (g/L)	129.34±14.48	116.12±20.23	<0.001
HCT (%)	38.86±3.98	35.13±5.98	<0.001
MCV (fL)	86.48±7.11	88.10±9.91	0.026
MCH (pg)	28.68±2.59	28.96±4.01	0.261
MCHC (g/L)	331.46±13.48	329.28±17.27	0.077
RDW-CV(%)	13.88±2.88	15.92±5.90	<0.001
RDW-SD (fL)	46.28±19.97	49.86±10.58	<0.001
Platelet count (10 ⁹ /L)	139.34±75.92	87.04±49.59	<0.001
MPV (fL)	11.54±1.59	11.16±1.53	0.009
PDW (fL)	17.05±15.96	15.52±1.95	0.013
PCT (%)	0.17±0.07	0.12±0.05	<0.001
P-LCC (10 ⁹ /L)	49.88±23.76	30.83±18.06	<0.001
P-LCR (%)	37.78±9.24	36.47±9.98	0.094
Urinary copper (μmol/L)	3.72±1.78	4.55±2.82	0.003
Serum ceruloplasmin (mg/L)	52.48±68.63	56.94±40.44	0.039

HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW-CV, red cell distribution width CV; RDW-SD, red cell distribution width SD; MPV, mean platelet volume; PDW, platelet distribution width; PCT, plateletcrit; P-LCC, platelet large cell count; P-LCR, platelet large cell ratio; NLR, neutrophil to lymphocyte ratio.

On the basis of Gini impurity, the predictors of importance in the model for prediction of liver cirrhosis are summarized in Fig. 2. The most important predictors in the predictive model were P-LCC, RDW-CV, serum ceruloplasmin, age at diagnosis, and MCV. But the relative

importance of usage of zinc and usage of DMSA/DMPS is equal to zero.

Next, we performed XGBoost with the same variables, and we found that the predictive model achieved an AUC of 0.9998, an accuracy of 0.996 [95% confidence interval

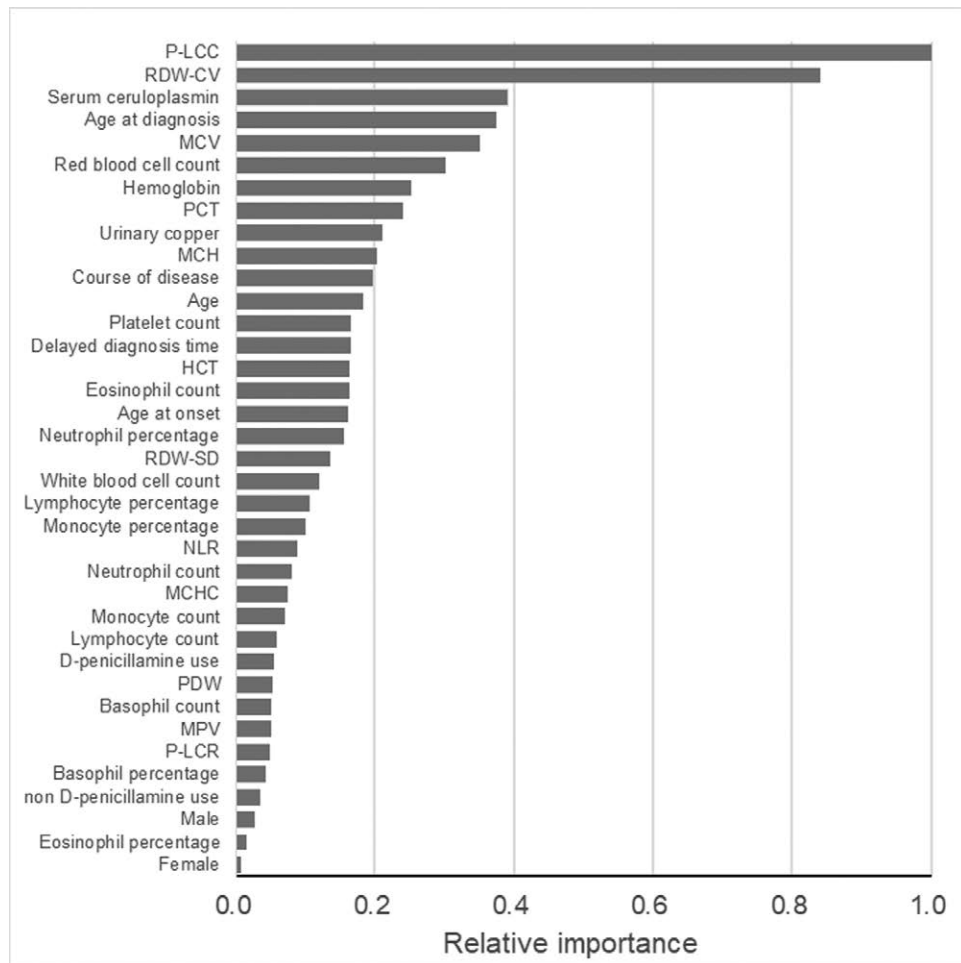


Fig. 2. Relative importance of the predictor variables in the predictive model. The relative importance score of predictor variables is based on Gini impurity. The most important predictor variable was set as 1, and the score of other variables is the ratio of the variable to the important variable. *P-LCC*, platelet large cell count; RDW-CV, red cell distribution width CV; MCV, mean corpuscular volume; PCT, plateletcrit; MCH, mean corpuscular hemoglobin; HCT, hematocrit; RDW-SD, red cell distribution width SD; NLR, neutrophil to lymphocyte ratio; MCHC, mean corpuscular hemoglobin concentration; PDW, platelet distribution width; MPV, mean platelet volume; *P-LCR*, platelet large cell ratio.

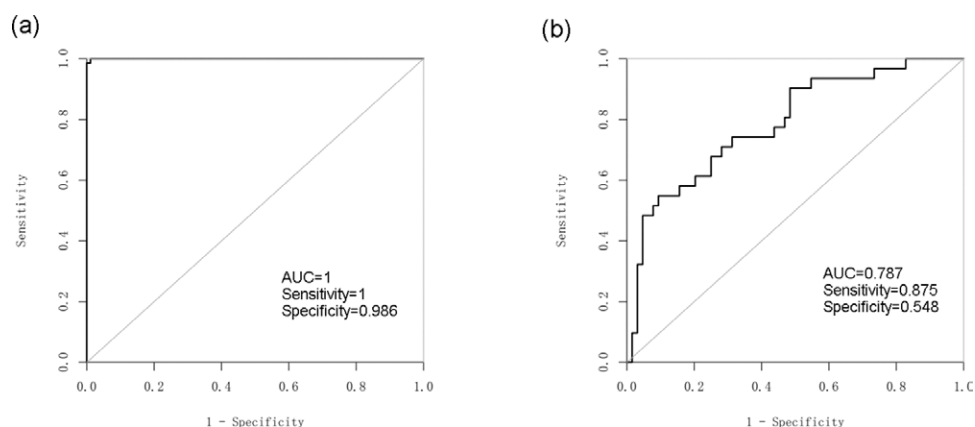


Fig. 3. Receiver operating characteristic curve (ROC) of predictive model. The vertical coordinate indicates true-positive rate (TPR, sensitivity), the horizontal coordinate indicates false-positive rate (FPR, 1-specificity) in (a) (training set) and (b) (testing set). ROC, receiver operating characteristic curve; AUC, area under the ROC; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW-CV, red cell distribution width CV; RDW-SD, red cell distribution width SD; MPV, mean platelet volume; PDW, platelet distribution width; PCT, plateletcrit; *P-LCC*, platelet large cell count; *P-LCR*, platelet large cell ratio; NLR, neutrophil to lymphocyte ratio.

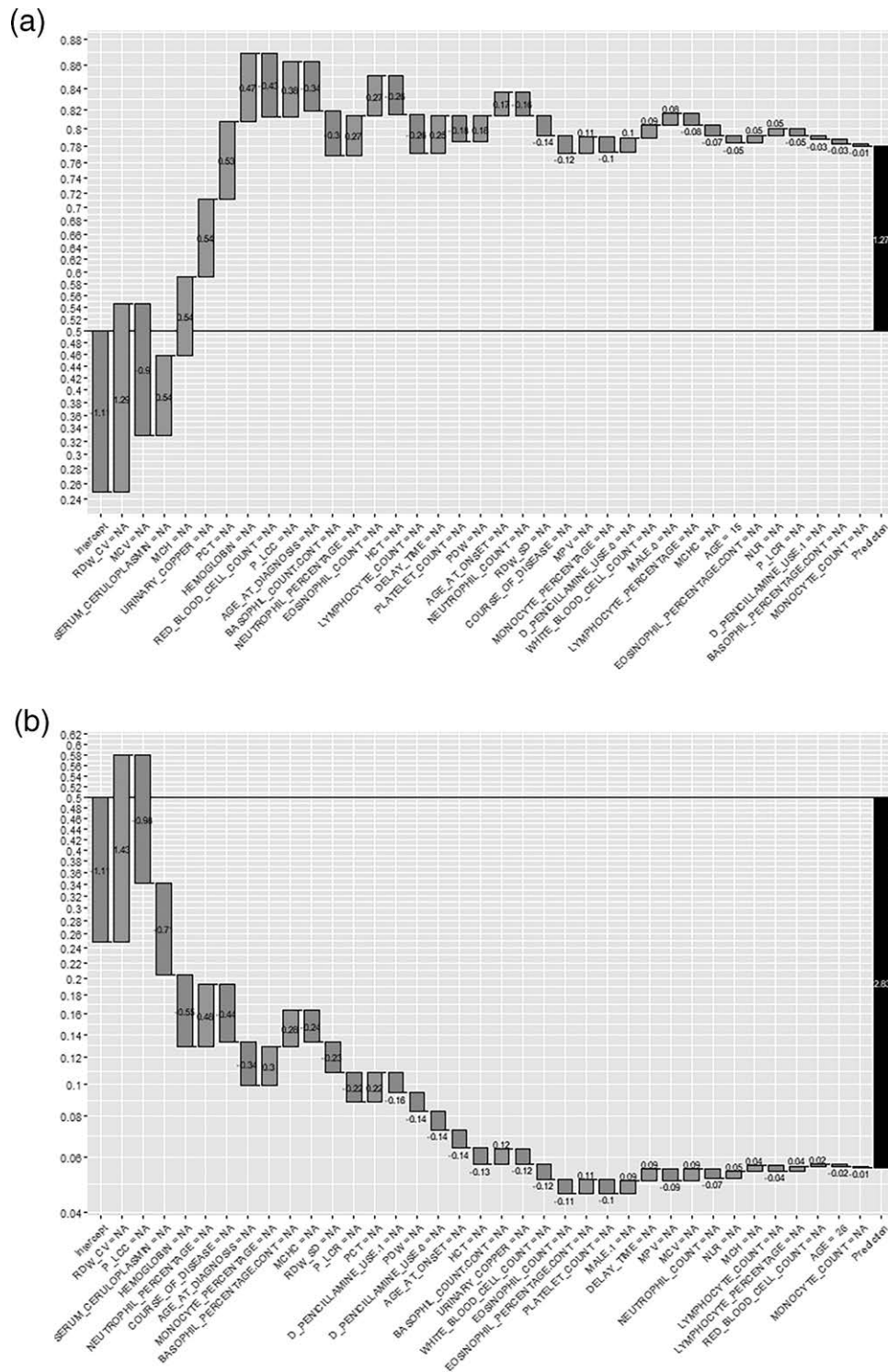


Fig. 4. The individualized predicted probability of study subjects. The vertical coordinate indicates predictive variables, the horizontal coordinate indicates the predicted probability of liver cirrhosis. The box in red means the variable leads to a decreased probability of occurrence of liver cirrhosis. The box in green means the variable leads to an increased probability of occurrence of liver cirrhosis. And the number in box means the relative importance of variables. The predicted probability of (a) is more than 0.5, and the predicted probability of (b) is less than 0.5. ROC, receiver operating characteristic curve; AUC, area under the ROC; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW-CV, red cell distribution width CV; RDW-SD, red cell distribution width SD; MPV, mean platelet volume; PDW, platelet distribution width; PCT, plateletcrit; P-LCC, platelet large cell count; P-LCR, platelet large cell ratio; NLR, neutrophil to lymphocyte ratio.

(CI), 0.9780–0.9999] in the training set. In the testing set, the AUC was 0.7873, with an accuracy of 0.7684 [95% CI, 0.6706–0.8488]. The results are shown in Fig. 3. As shown in Fig. 4, the individualized predicted probability of more than 0.5 was shown in Fig. 4a, and less than

0.5 was shown in Fig. 4b. As shown in Fig. 4a, variables (included age, age at diagnosis, course of disease, usage of D-penicillamine, neutrophil count, lymphocyte count, monocyte count, basophil count, lymphocyte percentage, monocyte percentage, basophil percentage, red blood cell

count, HCT, MCV, MCHC, RDW-SD, platelet count, and P-LCR) led to a decrease of probability of occurrence of liver cirrhosis. However, other variables (included female, age at onset, delayed diagnosis time, serum ceruloplasmin, urinary copper, white blood cell count, eosinophil count, NLR, neutrophil percentage, eosinophil percentage, hemoglobin, MCH, RDW-CV, MPV, PDW, PCT, and P-LCC) led to an increase of probability of occurrence of liver cirrhosis.

Discussion

In this retrospective cohort study, we applied machine learning to construct a predictive model for predicting the occurrence of liver cirrhosis in patients with WD. And we found P-LCC, RDW-CV, serum ceruloplasmin, age at diagnosis, and MCV were the top five important predictors according to the Gini impurity. Moreover, the model was of high accuracy, with an AUC of 0.9998 in the training set and 0.7873 in the testing set.

A previous study found that the common mutation was ATP7B c.813C > A mutation in patients with WD in the eastern part of India [13]. Next, they utilized the random forest to conduct a predictive model for predicting the common mutation in patients with WD, and they found that the most important features were gait, age at diagnosis, and dystonia [13]. Moreover, the predictive model is moderate, with an accuracy of 0.58 [13]. Wang et al. found that liver stiffness evaluated by two-dimensional real-time shear wave elastography, whose inflection point was 10.45 kPa, could predict the occurrence of hypersplenism [14]. Another study found that acoustic radiation force impulse and fibrosis-4 could be used to distinguish cirrhosis in patients with WD, which could partly replace liver biopsy [15]. Some prognostic scoring systems, including the MELD score and Nazer score, have been developed to identify poor prognosis in patients with WD with acute liver failure [16]. A prognostic system developed by the Acute Liver Failure Study Group with the random forest method could predict the progression of acute liver failure in patients with acute liver injury [17]. In the present study, we used general clinical data and blood routine examination to construct a model for predicting the development of cirrhosis in patients with WD. And the model was of high accuracy, with an AUC of 0.9998 in the training set and 0.7873 in the testing set.

Blood-based peripheral blood parameters have predictive and prognostic value in liver cirrhosis. Patients with autoimmune hepatitis had higher RDW-CV and RDW-SD value, as compared with healthy control [18,19]. And RDW-CV and RDW-SD could distinguish liver cirrhosis from chronic hepatitis B (CHB) and inactive hepatitis B virus (HBV) carriers [20]. In addition, RDW-CV is a valuable predictor for predicting survival and length of hospitalization in patients with liver cirrhosis [21]. And CHB with RDW-CV more than 15.1% had a higher mortality [22]. Recently, some ratios according to the peripheral blood parameters are available prognostic indicators for a liver disorder. The hemoglobin to RDW-SD ratio had a value to predict the survival in HBV-related decompensated cirrhosis [23]. Meanwhile, RDW-CV to platelet ratio could predict fibrosis in patients with chronic hepatitis C, with an AUC of 0.65 [24]. And the RDW-CV to

lymphocyte ratio could predict the incidence of CHB liver cirrhosis [25]. Previous study had found that the alcoholic cirrhosis group had a higher level of MCV value compared with alcoholic fatty liver or mild alcoholic hepatitis [26]. In addition, another study had found that higher MCV value is associated with the severity of HBV-related decompensated cirrhosis [27]. In the present study, we found that RDW-CV and MCV are the most important indices to predict the incidence of liver cirrhosis in WD. These results were consistent with previous studies.

The limitations of the study were as follows. Firstly, the study lacks external validation, and there is a risk of overfitting. Secondly, the study was a retrospective study, which meant selection bias and recalls bias existed. Finally, the prediction model constructed by machine learning is a black box, which couldn't present the weight of each variable. Therefore, a multicenter and prospective study and a web-based release prediction model should be conducted to settle these problems in the future.

Conclusion

In conclusion, the predictive model for predicting liver cirrhosis in WD, constructed by machine learning, had a higher accuracy. And the most important indices in the predictive model were P-LCC, RDW-CV, serum ceruloplasmin, age at diagnosis, and MCV.

Acknowledgements

H.P. proposed the idea; H.P., C.K., W.Y., M.J., L.Y., and Z.-m.G. acquired the data. H.P., C.K., and W.Y. analyzed the data. H.P. wrote the first draft; All authors have approved the final article.

The datasets of the current study are available from the corresponding author on reasonable request.

This study was approved by West China Fourth Hospital Ethics Committee (Approval No.: HXSJ-EC-2020088) and was performed in line with the Declaration of Helsinki.

Conflicts of interest

There are no conflicts of interest.

References

- Lo C, Bandmann O. Epidemiology and introduction to the clinical presentation of Wilson disease. *Handb Clin Neurol* 2017; 142:7–17.
- Chang J, Hahn SH. The genetics of Wilson disease. *Handb Clin Neurol* 2017; 142:19–34.
- Scheiber IF, Brüha R, Dušek P. Pathogenesis of Wilson disease. *Handb Clin Neurol* 2017; 142:43–55.
- Boga S, Ala A, Schilsky ML. Hepatic features of Wilson disease. *Handb Clin Neurol* 2017; 142:91–99.
- Asrani SK, Kamath PS. Model for end-stage liver disease score and MELD exceptions: 15 years later. *Hepatol Int* 2015; 9:346–354.
- Pogorzelska K, Krętowska A, Krawczuk-Rybak M, Sawicka-Żukowska M. Characteristics of platelet indices and their prognostic significance in selected medical condition - a systematic review. *Adv Med Sci* 2020; 65:310–315.
- Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Crit Care Med* 2016; 44:368–374.
- Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, et al. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response

- to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Invest Radiol* 2019; 54:110–117.
- 9 European Association for Study of Liver. EASL clinical practice guidelines: Wilson's disease. *J Hepatol* 2012; 56:671–685.
 - 10 Saverymuttu SH, Joseph AE, Maxwell JD. Ultrasound scanning in the detection of hepatic fibrosis and steatosis. *Br Med J (Clin Res Ed)* 1986; 292:13–15.
 - 11 Ding XF, Li JB, Liang HY, Wang ZY, Jiao TT, Liu Z, *et al.* Predictive model for acute respiratory distress syndrome events in ICU patients in China using machine learning algorithms: a secondary analysis of a cohort study. *J Transl Med* 2019; 17:326.
 - 12 Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw* 1999; 10:988–999.
 - 13 Chaudhuri J, Biswas S, Gangopadhyay G, Biswas T, Datta J, Biswas A, *et al.* Correlation of ATP7B gene mutations with clinical phenotype and radiological features in Indian Wilson disease patients. *Acta Neurol Belg* 2022; 122:181–190.
 - 14 Wang J, Hu M, Zhu Q, Sun L. Liver stiffness assessed by real-time two-dimensional shear wave elastography predicts hypersplenism in patients with Wilson's disease: a prospective study. *BMC Med Imaging* 2022; 22:25.
 - 15 Li Y, Ma J, Li B, Zhu X, Wang J. Cirrhosis of Wilson's disease: high and low cutoff using acoustic radiation force impulse (ARFI) -Comparison and combination with serum fibrosis index. *Clin Hemorheol Microcirc* 2021; 79:575–585.
 - 16 Camarata MA, Gottfried M, Rule JA, Ala A, Lee WM, Todd Stravitz R, Schilsky ML. Outcomes of acute liver injury in adults due to Wilson's disease: is survival without transplant possible? *Liver Transpl* 2020; 26:330–336.
 - 17 Koch DG, Speiser JL, Durkalski V, Fontana RJ, Davern T, McGuire B, *et al.* The natural history of severe acute liver injury. *Am J Gastroenterol* 2017; 112:1389–1396.
 - 18 Liu L, Cao J, Zhong Z, Guo Z, Jiang Y, Bai Y, Xu J. Noninvasive indicators predict advanced liver fibrosis in autoimmune hepatitis patients. *J Clin Lab Anal* 2019; 33:e22922.
 - 19 Wang H, Xu H, Qu L, Wang X, Wu R, Gao X, *et al.* Red blood cell distribution width and globulin, noninvasive indicators of fibrosis and inflammation in chronic hepatitis patients. *Eur J Gastroenterol Hepatol* 2016; 28:997–1002.
 - 20 Zhu M, Han M, Xiao X, Lu S, Guan Z, Song Y, Liu C. Dynamic differences of red cell distribution width levels contribute to the differential diagnosis of hepatitis B virus-related chronic liver diseases: a case-control study. *Int J Med Sci* 2019; 16:720–728.
 - 21 Abdel-Samiee M, Ahmed Samy Kohla M, Fathy Ragab M, Ghanem HS, Kamel AbouGabal A, Abdelsameea E. Red cell distribution width as a predictor of outcome in hospitalized cirrhotic patients. *Eur J Gastroenterol Hepatol* 2021; 33:e978–e985.
 - 22 Wang J, Huang R, Yan X, Li M, Chen Y, Xia J, *et al.* Red blood cell distribution width: a promising index for evaluating the severity and long-term prognosis of hepatitis B virus-related diseases. *Dig Liver Dis* 2020; 52:440–446.
 - 23 Yu Z, Zhang T, Shen J. Low hemoglobin-to-red cell distribution width ratio is associated with mortality in patients with HBV-related decompensated cirrhosis. *Biomed Res Int* 2022; 2022:5754790.
 - 24 Ramzy I, Fouad R, Salama R, Abdellatif Z, Elsharkawy A, Zayed N, *et al.* Evaluation of red cell distribution width to platelet ratio as a novel non-invasive index for predicting hepatic fibrosis in patients with chronic hepatitis C. *Arab J Gastroenterol* 2021; 22:6–11.
 - 25 Zhang X, Wang D, Chen Z, Guo N, Wang W, Xiong C, *et al.* Red cell distribution width-to-lymphocyte ratio: a novel predictor for HBV-related liver cirrhosis. *Medicine (Baltimore)* 2020; 99:e20638.
 - 26 Huang A, Chang B, Sun Y, Lin H, Li B, Teng G, Zou ZS. Disease spectrum of alcoholic liver disease in Beijing 302 Hospital from 2002 to 2013: A large tertiary referral hospital experience from 7422 patients. *Medicine* 2017; 96:e6163.
 - 27 Yang J, Yan B, Yang L, Li H, Fan Y, Zhu F, *et al.* Macrocytic anemia is associated with the severity of liver impairment in patients with hepatitis B virus-related decompensated cirrhosis: a retrospective cross-sectional study. *BMC Gastroenterol* 2018; 18:161.